Dream team
Methodology

Technology

Science
Approach

1) **No ready solutions** → perfect customisation
   (our code)

2) **Modular development** → keep time and costs under control
   (OOP Objects Oriented Programming)

3) **Abstract approach** → keep development under control
   (UML Unified Modelling Language design)

4) **Lateral thinking** for innovation → “Everyone has always done in this way” is not necessary a value
   (high regular investments in R&D)

5) **Continue updating** → To be always on the bleeding edge of the technology
   (35% R&D investment)
20 years of innovation (a)

Starting from the past millennium several innovative jobs:

- **1999 Live Streaming** with Nilla Pizzi during Sanremo Festival
- **2000 Live Streaming of the first version of Italian “Big Brother”**
- **2000 Nixie: one of the first e-commerce case of success in Italy grown up till the creation of physical shops (followed by many e-commerce sites such as the actual: sportit and dimanoinmano)**
- **2001 e-learning Interactive solution for Sanofi (Pharma company) (slide sync with streaming and tests)**
- **2001 A very automated network of 54 sites of the 54 towns of the future Monza Province, with events, appointments and databases of useful information**
- **2002 Complete automated online Catalogue with CD syncing for Selva (the biggest furniture company in Italy)**
- **2002 One of the first editorial online magazines for EditriceTemi (followed by many other success cases with EditorialeDirecta, Sunnycom, EditorialeDomus, Radiolombardia, Accordo.it, Quattroruote ...)**
- **2002 A flash screensaver for all ST Microelettronics PCs in the world to celebrate their 15th year.**
- **2002 One of the first social community in Italy for Radiolombardia (Discoletto) (4 years before Facebook)**
- **2002 Online catalogue for the photos (RM and RF) for Olycom (the biggest Italian photos agency)**
- **2003 A very modern site for the 10 cinema in Monza, allied to create a “virtual” multiplex to fight real multiplex. Programming, automated newsletter, booking and purchasing of ticket**
- **2003 A Windows-based ERP for Banca Popolare di Milano (CRAL) – managing 40M € / year, selling everything from food, to vacations, from jewels to consultant.**
- **2004 Georefering Selva’s shops in all Europe with map and information to get to the nearest one (2 years, before Google)**
- **2005 In browser 3D projects solutions for Rifra in Flash technologies (ActionScript)**
- **2005 Management software for a Clinical Center**
- **2006 Happyrunner: A virtual runners' and triathletes' society that soon became one of first five in Italy (currently more then 1.300 members).**
- **2007 Interactive flip system to consult research results of Roche for their Intranet**
- **2008 Innovative intranet for Havi, international logistic company (supplier also of McDonald's all over Europe)**
20 years of innovation (b)

- **2009** Module to interact with Regione Lombardia **certified by SISS** (smartcard for public health)
- **2010** One of the first **radio app** in Italy (live with titles in streaming and on demand) for Radiolombardia,
- **2010** powerful **crawler** for fashion (collecting articles and advertisement)
- **2011** One of the first live **assistant with videochat** in the Tisettanta outlet website
- **2012** Innovation of accordo.it getting the first web community in Italy (30.000 users)
- **2013** Forecasting solutions for McDonalds’
- **2014** **Big Data Solutions** for Sky (billion of records with thousands of features) in support of Adobe Systems.
- **2014** The first **industrial app** in the world to manage big production machine **via modbus** (refused at the first time by Apple because they didn’t even understand the scope!)
- **2014** A revolutionary **3D app** that reproduce a physical game of Geomag
- **2015** **Real time advertising solution** for Sky (+400% performance)
- **2016** A powerful search engine for e-commerce sites, with similarity and automated filters
- **2016** **Antichurn** web application for Italtel presented at MWC
- **2016** An innovative **messenger** which represents the mood of users in the balloon
- **2016** **Recommendation** engine for editorial site for Deutsche Welle AppleTV app.
- **2017** Collection and analysis of the advertising online for JCDecaux
- **2017** Recommendation engine for e-commerce for sportit.com
- **2017** Analysis of the web to map customer satisfaction of brand and its competitors
- **2018** **CRM Enrichment** for Banca Intesa
- **2018** **Hand writing recognition** for Generali
- **2018** **Meta-sotware** for McDonald's and Aspis
- **2019** **Unstructured data discovery** for Essilor
- **2019** Recommendation engine for the Marketing of GEDI (Gruppo Espresso)
- **2019** **Smart living** project with Evolvere
- **2019** **Industria 4.0** project with Italbond
But also a lot of business as usual

- hundreds of **customers**
- hundreds of **websites**
- tens of **apps**
- tens of management **software**
- tens of **infrastructures** managed
- hundreds of **hostings**
- tens of **housings**
- ...

"Sinte"
Infrastructure

- 30 physical superservers in different webfarms around Europe
- More than 150 Virtual Servers
- More than 400 managed domains
- More than 5,000 mailboxes on our mailserver
- More than 500 web application developed
Competences

- Software
- Web (e-commerce, e-learning, streaming...)
- Mobile apps
- Data collection
- Security
- Image analysis
- IT infrastructures
- Hosting & housing
- System integration
- Statistics
- Big data solutions
- Data science
- Data visualization
- Semantic
- Machine learning and deep learning
- Artificial Intelligence
- Usability
smartstat : Sinte’s spin-off
Business Science
for governance, competition and innovation
www.smartstat.it
Linking two wor(l)ds to share value

ACADEMY ↔ MARKET
Mission

- Fill the gap between the big amount of meaningful data and the power of methodology’s vanguard
- Fill the gap between the data and the information needed
- Use data:
  - to forecast,
  - to deeply know your own market,
  - to intercept signals of changes ...
- To provide companies the powerful tools to guarantee their future prosperity
The duty to innovate

“That's the way it's always been done” is often the reason to behave in a new way

Quite always solutions:

● Are not written in manual
● Are not stored in paradigms
● Cannot be found in the past

We all have to:

● Think in a new way
● Look for correlations
● Use imagination
● Focus on continue confrontation with academic research
What we use

- **Machine learning**: all processes are developed in solution able to continuously learn from the data flow to:
  - improve performance
  - identify seasonality
  - adapt to the continues changes

- **Artificial Intelligence**: our systems are a complex structure of software solutions capable to simulate human ability to learn from reality and to adapt to it

- **Embedding of human intelligence** know-how: use of Bayesian models to allow human to teach the engine how to better do its job

- **Neural Network**: Simulating the approach of human brain in some situation system are able to achieve impressive results in term of performance and delivery time

- **Deep learning**: with the heavy use of neural network we can easily develop powerful deep learning solution

- **Vanguard sophisticated models**: we usually develop solutions using all the bleeding edge model of the international statistics research reading the source code and modifying them according to our needs (for instance we have used genetics biostatistical models adapted to analyze cookies)
Innovative approaches

• Beyond one-size-fits-all solutions -> **ad hoc solution**
  Agnostic and continuously updated approach towards:
  - technologies
  - statistical models
• Beyond just big data -> **complex data**
  Open mind approach using not "superclean" data and looking for innovative ideas to improve their informational power.
  Crucial step is data assessment to feed models with meaningful data – to do this most innovative and sophisticated models and approach are needed (if not we get “garbage in - garbage out”)
• Beyond “data science” -> **Business Science**
  Approach situation as a real scientists (both founders are graduated in Physics)
• Beyond “analytics” -> **stylization**
  Always towards new way to see the data according to the sensibility, the time and the needs of management
• Beyond “look alike” -> **functional profiling**
  We started from the beginning approaching our project with functional profiling
Recursive approach

data collection

lateral thinking

data visualization

business practice

stylistization propension

scientific approach

data management

data elaboration

Business practice and scientific approach are closely intertwined in this recursive process, as they form the foundation for lateral thinking and data visualization. This cycle allows for continuous refinement and innovation in the field of data management.
Features quest

RAW DATA

- CLASSIFIERS
- SEMANTIC
- PROCESSING
- STATISTICAL

human insights

NEW PROFILE

STATISTICAL CHECK

SUPER MEANINGFUL PROFILE
The only Statistics University in Lombardia
Advantages

• Work side-by-side with many researcher and professors in different specialisation
• A network of researchers from all the world
• Contact with all the top students in data-science in Italy
Bicocca University departments

- Economia, metodi quantitativi e strategie di Impresa
- Scienze economico-aziendali e diritto per l'economia
- Statistica e metodi quantitativi
- Giurisprudenza
- Medicina e chirurgia
- Psicologia
- Biotecnologie e bioscienze
- Fisica "Giuseppe Occhialini"
- Informatica, sistemistica e comunicazione
- Matematica e applicazioni
- Scienza dei materiali
- Scienze dell'ambiente e della Terra
- Scienze umane per la formazione "Riccardo Massa"
- Sociologia e ricerca sociale
Some cases
Restaurants management

- **Request**: A easy-to-use software to manage a multi-restaurants company at-a-glance.
- **Approach**: Create a very customisable software able to adapt itself to every need of the different franchisers.
- **Results**: A very appreciated system used now in 45 McDonald’s restaurants (number increasing)
Price algorithm

- **Request:** To automatically identify the right price for an order of metal bars or plates of the dimensions requested by the customer in an e-commerce site.

- **Approach:** A powerful collecting engine that gathers prices of base metal and alloys from different sources (such as London Metal Exchange and Assomet database) feeds a sophisticated algorithm that considers tens of variables (such as time to cut the metal, chippings of the cut, waste parts, the need to use fork lift, the stock amount, the prices timeseries) to determine the correct sell price.

- **Results:** An engine in production from 2005 (14 years).
Mobile app for industrial production

- **Request**: To develop a mobile app able to interact with industrial 24h-functioning machines

- **Approach**: a cross-platform mobile app that allows to view and modify values in the machine and a web application that communicates both with the machines (via modbus) and with the mobile app

- **Results**: A solution that allows production firm to control all the machines from wherever and to receive push notification according to settings customizable for each machine.
3D mobile app for game simulation

- **Request**: Create an app able to simulate the real game in 3D
- **Approach**: An hard work to import or the 3D models of every pieces to reduce computational effort of the device maintaining the image quality
- **Results**: The very first app of this genre in the world
Industria 4.0 solution

- **Request**: To create a solution able to monitor all the production phases of the paper enrichment.

- **Approach**: To develop an open solution to control all the production using Arduino’s devices and Raspberry mini PCs in order to detect and record different parameters of each machine.

- **Results**: The firm is now able to control all the processing.
Handwriting recognition

- **Request**: Automatically identify the price and the date from the images of taxi receipts sent by employers using their cellphones.

- **Approach**: A mix of innovative pattern recognition solution, smart algorithms and machine learning systems to deliver the result.

- **Results**: the first engine identify the correct data in more than 30% of the receipts (where the former suppliers arrived to 4%, 7,5 times better than our competitor).
Image pattern recognition

- **Request**: Automatically identify the correct image taken in different angle, surface and illumination.

- **Approach**: A pattern recognition engine able to identify images in every condition.

- **Results**: the first engine identify the correct image in 92% of cases.
IAB categories click optimisation - 1

- **Request**: increase the click rate on banners on top of the Google optimisation, having information of just 30% of traffic of campaigns, dividing users in 5 clusters for each of the first level of IAB categorization (customer is not able to perform a one-to-one delivery).

- **Approach**: create high significant profile of each cookie using semantic and classifications and create sophisticate models to predict the click rate of each user for each category.

- **Results**: Increase in quite all categories, and in the most focused one the click rate is **300%** the former one.
IAB categories click optimisation - 2

Marzo

0% 2% 4% 6% 8% 10% 12%

4 3 2 1 0

Sinte
Campaign click optimisation

- **Request**: increase the click rate on banners for their campaign.

- **Approach**: create high significant profile of each cookie using semantic and classifications and create sophisticate models to predict the click rate of each user for each category.

- **Results**: The result of our cookies are between **300% and 700%** our competitor’s results.
Antichurn - 1

- **Request**: predict the churn probability of each customer just according to their browsing behaviour outside our clients’s website (wherever are its banners). No info on their socio-demo, their contract. In this very first step, no distinction between recent and distant browsing behaviour.

- **Approach**: A machine learning system feeded with a very smart profiling of users according to the needs.

- **Results**: the last quantile has **400%** the users of the first one. Now it is ready to regularly enrich the CRM system with new KPI.
Home page sorting

- **Request**: Create a solution to differentiate the order of offers in the homepage to increase sells. For technological limits of the customer we cannot suggest products and we cannot create a real time approach one-to-one, but we have to work to create the sorting every night for the day after clustering users in 35 clusters.

- **Approach**: Using an intelligent mix of ordinal variables models and cardinal variables models we create a solution in machine learning which will increase performances.

- **Results**: The result is +8% in incomes in the first month.
Editorial recommendation

- **Request**: Create a propensity model to populate a box of “chosen for you”.

- **Approach**: We use three engines. A themes collector which gathers tendencies and trends from the web, a semantic engine which define very precise profile of contents and a statistical engine which provide propensity models.

- **Results**: **Over 2 times better then former engine**
Marketing recommendation

• **Request:** Create a propensity model to help Marketing to identify the best target of a campaign.

• **Approach:** We used all data gathered from all the properties of our client to create a profile of thousands of features.

• **Results:** A solution integrated in the marketing UI to help Marketing suggesting solution but also capable to deliver the best optimisation without any human decision.
E-commerce recommendation

- **Request**: Create different engines to propose the best set of products for each single user in different part of the site:
  - home page “most suitable for you”, “similar to previously viewed products”, “offers for you”, “most sold”
  - product page: “alternatives” and “to buy together”

- **Approach**: a mix of statistical models and very customized algorithm with a deep knowledge of e-commerce logics.

- **Results**: *Over 6 times better than former engine.*
• **Request:** To find a way to use Nielsen’s Data of investment in advertising of Italian companies.

• **Approach:** Using our customisation of Kohonen maps (neural maps) we realise a meaningful representation of a multidimensional environment in a human readable two dimensional map.

• **Results:** *Innovative solution* which delivers regularly the maps used constantly by the management to decide strategies.
Marketing map - 2

Automotive

Automobile

Magazines

Newspapers

Fashion

Abbigliamento

Out of home

Internet
Prospecting

• **Request**: To substitute former systems which give a score of probability to acquire each prospect.

• **Approach**: Using ordinal variables models (typical of social science) we developed a revolutionary solution that:
  – asks to human inputs more similar to their way to think
  – creates a score mathematically correct

• **Results**: a new model able to identify company that will make the 90% in a group **5 times more narrow** than the former model.
Restaurants forecasting

- **Request**: A system able to continuously predict the revenues in the next two weeks

- **Approach**: Using advanced time series models and giving a lot of external features (holidays, weather, sales periods, Lent, Ramadan, soccer matches ...) we realize a solution able to give trend and forecasting of the revenues per each days in the future.

- **Results**: forecast the revenues of each restaurant with **median absolute error of 4.19%** in the last (the 14th) day in the future.
E-commerce forecasting - 1

- **Request**: A system able to predict sell in the next months divided by type of product for each brand’s shop.

- **Approach**: Embedding our client’s know-how and external events in time series models we realize a solution able to give trend and forecasting of the sells.

- **Results**: Three customizable engines (from the simplest to the more complex) able to forecast the revenues of each brand just changing the data feed.
E-commerce forecasting - 2

Regressors:
S=Sales
C=Calendar
F=Festivity
A=Autoregression

Sinte
Sentiment in political themes

- **Request**: An observatory for opinion on social network about political subjects

- **Approach**: Using a mix of text mining, semantic and sentiment algorithm and high customised mathematical model we create an engine which analyse what people says in Twitter about themes given in input by our customer

- **Results**: a solution which gives the ratio between positive and negative sentiment on every theme the politicians are interested in.
Search engine optimisation

- **Request:** A powerful search engine which could give in real-time the results in order of matching (using also similarity algorithms) and all possible filters at the same time.

- **Approach:** An advanced mix of full text and structured data search solution together with high optimisation of indexes and auto-learning models.

- **Results:** +22% page viewed per session
Web crawling

- **Request**: Get unstructured informations from web (articles, posts, advertising, photos...) transforming them in structured data for analysis.

- **Approach**: Create a very intelligent crawler capable of acting and “thinking” as human.

- **Result**: An engine that collect **4 times articles** than the best competitor and collect **advertisement where no one else succeeded**.
Listening of the Web - 1

- **Request**: Detect and measure customers’ perception of the service from Internet.

- **Approach**: Using a very powerful crawling engine, human trained semantic and sentiment algorithm and innovative mathematical model developed by us we create an engine which analyse what people says in Internet and social networks identifying the type of problem and where it is.

- **Results**: a solution which gives an hourly feed of information about: the problem, the devices impacted, the geolocation and the environment. Tested in production it detect and correctly identify in real time more the 90% of the problems.
Listening of the Web - 2
Meta analysis

- **Request**: Find a new scientific way to measure effectiveness, efficacy and safety of drugs and medical devices already in the market, without the big cost of the clinical research.

- **Approach**: Using the big amount of data in scientific research from publications (which aren’t focused on this scope) and elaborate them with innovative models able to find scientific measurement deduced from the intersection of this apparently inhomogeneous data.

- **Result**: Strong and **scientific results** with a precision comparable with the clinic research one **in less than 1/100 of costs and less than 1/10 of time**.
Sales-force evaluation

- **Request:** Find a way to compare performance of hundreds of business agents engaged in different areas, different clients (Carrefour, Auchan...) and different part of the catalogue.

- **Approach:** Study a new model that scientifically balance advantages and disadvantages according to data, “playing all agents in the same field”.

- **Result:** A new indispensable tool: a map with recalculated absolute and relative margin where our customer could really compare the performance.
Advertising observatory - 1

- **Request**: Get a map of advertising online to understand market changes and to intercept clients who are investing in advertising.

- **Approach**: Use our crawler to constantly monitor the advertising on the web detecting websites, page level, advertiser, product, banner position and frequency, giving a daily report.

- **Result**: An engine that collects all investors in advertising both in (inter)national sites and in local sites (about what there aren’t even aggregated informations)
Advertising observatory - 2

Corriere della sera

Dazn

Sinte
Socio-demo segmentation

- **Request**: Identify gender and age of anonymous cookies
- **Approach**: Use the deterministic data of customers to create a powerful model that is able to determine the age and gender of anonymous prospects according to their navigation profile.
- **First result**: With the very first model we got very good results: dividing all the cookies in males and females, in every group there are double the number of the other gender.
Market Analysis from cookies - 1

- **Request**: Use cookies’ anonymous and aggregated data to create valuable information.

- **Approach**: Use the big amount of information collected by cookie analysis to identify how particular target (ex: people interested in cosmetics) differs from the mean of population,

- **First result**: With a very smart approach we can get a map to show *how to act for every kind of target*. 
Customers clustering - 1

- **Request**: Identify the main clusters in the customer base using their behaviour on TV
- **Approach**: Create profile of hundreds of features for each customer and after a deep analysis of the correlations we use a centroid-based clustering.
- **Result**: We identified a easy process that allows to decide how many clusters and which features to use and it creates best clusters, gives different intelligent measures to evaluate the result and allows to navigate each cluster showing how it differs in the other non-selected features.
Customers clustering - 2

- Film: 1.442872
- Fiction: 1.207254
- Unknown: 1.087285

Ratio

- Cartoon: 0.394086
- News: 0.521293
- Intrattenimento: 0.838841
- Cultura: 0.87331

Difference

- 0.074
- 0.020305
- 0.027435
- 0.011
- 0.061
- 0.041
- 0.000
Smart Living

- **Request**: A solution that takes data from thousands of probes and sensors and forecast temperature and energy saving solution

- **Approach**: A big data scalable infrastructure with modern models

- **Result**: A solution that, thanks to its innovative approach, received public financing of 140,000 euro.
Unstructured data discover

- **Request**: Identify the content of files in the company’s repository to discover sensible data for GDPR and map the file in a complete GDPR solution.

- **Approach**: Create a solution as a mix of document parsing, regular expressions, list of data, sophisticated algorithm.

- **Result**: We are able to identify the sensible data inside Office documents, pdf, text files, images, audio and video.
BIG DATA, BIG CONFUSION?

smartstat
Data science for your business